Discretization Methods

Juan Castellanos
European University Institute

April 3, 2023

Intro

- Consider an AR(1) process of the form

$$z_t = \rho z_{t-1} + \varepsilon_t \tag{1}$$

with stationary distribution $\mathcal{N}(0, \sigma_z^2)$ where $\sigma_z = \frac{\sigma_\varepsilon}{\sqrt{1-\rho^2}}$

- **Goal**: obtain a finite state Markov chain that generates the same population moments as the continuous process

- **Why useful?**
  * Solving Bellman equations typically involves taking expectations over next period values
  * If process is continuous, then one has to deal with integrals which are computationally costly
  * If discrete, the continuation value is just a weighted sum

What is a Markov chain?

- Let X be a finite set with n elements : $\{x_1, \ldots, x_n\}$
  * Think of X as the *exogenous* state space and x_i as the state values

- A Markov chain on X is a sequence of random variables $\{X_t\}$ that satisfy for any date t and any next period state $\tilde{x} \in X$

$$\mathbb{P}(X_{t+1} = \tilde{x} | X_t) = \mathbb{P}(X_{t+1} = \tilde{x} | X_t, X_{t-1}, X_{t-2}, \ldots) \tag{2}$$

  * In words, knowing the current state is enough to form expectations about future states.
  * This is the so called Markov property

- The dynamics of a Markov chain are fully determined by the set of probability values

$$P(x, \tilde{x}) := \mathbb{P}(X_{t+1} = \tilde{x} | X_t = x) \quad (x, \tilde{x} \in X) \tag{3}$$

  * $P(x, \tilde{x}) \geq 0$ is the probability of going from x to $\tilde{x}$ in one step
  * $P(x, \cdot)$ is the conditional distribution of X_{t+1} given $X_t = x$

How to simulate a Markov chain?

- Assume that you have a stochastic matrix P, that is a $n \times n$ matrix such that
  * Each element of P is non-negative
  * Each row of P sums to one
- Then, you can generate a Markov chain $\{X_t\}$ as follows:
  * Set a initial value or draw it from a know distribution. Call it x_0.
  * Since you know x_0, you can draw X_1 using the $\mathbb{P}(X_1|X_0 = x_0)$
  * Thus, in general you can draw X_{t+1} using $P(X_t, \cdot)$ for all $t = 0, 1, \ldots$
- Let's see an example using the following stochastic matrix

$$P := \begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix} \tag{4}$$

  * You can think of $1-\alpha$ as the probability of finding a job conditional on being unemployed and $1-\beta$ as the probability of losing a job conditional on being employed.

Tauchen's Method

1. Choose a value for the persistence $\rho \in (0, 1)$ and the standard deviation $\sigma_z > 0$

2. Set values for the hyper-parameters
   * n : number of potential realization of the process
   * m : number of standard deviations away from the the unconditional mean
   * Typical values : n=$\{5, 7, 9, 11, 15\}$, $m = \{2, 3, 4\}$

3. Set the bounds for the process

$$\bar{z} = m\sigma_z \tag{5}$$

$$\underline{z} = -m\sigma_z \tag{6}$$

Tauchen's Method

EUROPEAN UNIVERSITY INSTITUTE

4. Set $\{z_i\}_{i=1}^{n}$ such that:

$$z_i = \underline{z} + \frac{\bar{z} - \underline{z}}{n-1}(i-1) \tag{7}$$

   and construct mid-points $\{\tilde{z}\}_{i=1}^{n-1}$ which are given by:

$$\tilde{z}_i = \frac{z_{i+1} + z_i}{2} \tag{8}$$

5. The transition probability $p_{ij} \in P_{z,z'}$ (the probability of going to state z_j conditional on being in state z_i) is computed according to

$$p_{ij} = \Phi\left(\frac{\tilde{z}_j - \rho z_i}{\sigma_\varepsilon}\right) - \Phi\left(\frac{\tilde{z}_{j-1} - \rho z_i}{\sigma_\varepsilon}\right) \quad j = 2, 3, \ldots, n-1 \tag{9}$$

$$p_{i1} = \Phi\left(\frac{\tilde{z}_1 - \rho z_i}{\sigma_\varepsilon}\right) \tag{10}$$

$$p_{in} = 1 - \Phi\left(\frac{\tilde{z}_{n-1} - \rho z_i}{\sigma_\varepsilon}\right) \tag{11}$$

   where $\Phi(\cdot)$ denotes a CDF of the $\mathcal{N}(0,1)$

Transition probabilites: some details

- Expression (9) is obtained as follows
  * Let $d = z_{k+1} - z_k$ be the distance between two points in the vector of state values.
  * Then,

$$
\begin{aligned}
p_{i,j} &= \Pr\left\{z' = z_j \mid z = z_i\right\} \\
&= \Pr\left\{z_j - d/2 < z' \leq z_j + d/2 \mid z = z_i\right\} \\
&= \Pr\left\{z_j - d/2 < \rho z_i + \varepsilon \leq z_j + d/2\right\} \\
&= \Pr\left\{\frac{z_j + d/2 - \rho z_i}{\sigma_\varepsilon} < \frac{\varepsilon}{\sigma_\varepsilon} \leq \frac{z_j - d/2 - \rho z_i}{\sigma_\varepsilon}\right\} \\
&= \Phi\left(\frac{z_j + d/2 - \rho z_i}{\sigma_\varepsilon}\right) - \Phi\left(\frac{z_j - d/2 - \rho z_i}{\sigma_\varepsilon}\right)
\end{aligned} \tag{12}
$$

An example: going to the computer!

- Assume that we have a stochastic process

$$y_t = 0.85 y_{t-1} + \varepsilon \quad \text{with} \quad \mathcal{N}(0, 0.0095^2) \tag{13}$$

- We want to approximate it with a Markov chain with 5 and 9 points. We set $m = 3$.
- How does the approximation depend on number of potential realizations of the process?
  * We simulate the Markov chain for T = 10,000 periods
  * Are the sample moments close to the population ones?
  * What about the persistence of the process?

	$\hat{\mu}_y$	$\hat{\sigma}_y$	$\hat{\rho}$
n = 5	1.42e-5	0.012	0.881
n = 9	-0.0001	0.010	0.851

Rouwenhorst's Method

EUROPEAN UNIVERSITY INSTITUTE

1. Choose a value for the persistence $\rho \in (0, 1)$ and the standard deviation $\sigma_z > 0$

2. Set values for the hyper-parameters
   * n : number of potential realization of the process
   * λ : controls the upper and lower bound of the process
   * Typical value for $\lambda = \sqrt{n-1}\,\sigma_z$

3. Set the bounds for the process as follows

$$\bar{z} = \lambda \tag{14}$$

$$\underline{z} = -\lambda \tag{15}$$

4. Set $\{z_i\}_{i=1}^n$ such that:

$$z_i = \underline{z} + \frac{\bar{z} - \underline{z}}{n-1}(i-1) \tag{16}$$

EUROPEAN UNIVERSITY INSTITUTE

5. When $n = 2$, let P_2 be given by

$$P_2 = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} \tag{17}$$

   * p and q can be set to $p = q = \frac{1+\rho}{2}$

6. For $n > 2$, construct *recursively* the transition matrix:

$$P_n = p \begin{bmatrix} P_{n-1} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix} + (1-p) \begin{bmatrix} \mathbf{0} & P_{n-1} \\ 0 & \mathbf{0}' \end{bmatrix} + q \begin{bmatrix} \mathbf{0}' & 0 \\ P_{n-1} & \mathbf{0} \end{bmatrix} + (1-q) \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & P_{n-1} \end{bmatrix} \tag{18}$$

   where $\mathbf{0}$ is a $(n-1) \times 1$ column vector of zeros.

7. Divide all elements in the middle rows (except top and bottom) by 2 so the sum of each row is equal to 1. The final outcome is $P_{z,z'}$

An example: $n = 3$

- Let n = 3 and P_2 be given by equation (17). Then,

$$\begin{aligned} P_3 = & p \begin{bmatrix} p & 1-p & 0 \\ 1-q & q & 0 \\ 0 & 0 & 0 \end{bmatrix} + (1-p) \begin{bmatrix} 0 & p & 1-p \\ 0 & 1-q & q \\ 0 & 0 & 0 \end{bmatrix} \\ & + q \begin{bmatrix} 0 & 0 & 0 \\ p & 1-p & 0 \\ 1-q & q & 0 \end{bmatrix} + (1-q) \begin{bmatrix} 0 & 0 & 0 \\ 0 & p & 1-p \\ 0 & 1-q & q \end{bmatrix} \end{aligned} \tag{19}$$

- After multiplying and re-arranging terms we obtain

$$P_3 = \begin{bmatrix} p^2 & 2p(1-p) & (1-p)^2 \\ p & 1 & 1-p \\ (1-q)q & q^2 + (1-q)^2 & (1-q)q \end{bmatrix} \tag{20}$$

- Second row sums up to 2! Not consistent with definition of stochastic matrix ...

- That is why we divide by 2.

Approximations to highly persistent processes

- Kopecky and Suen (2010, RES) show that the Rouwenhorst method is superior when the process is highly persistent
- Assume we have stochastic process

$$x_t = 0.975 x_{t-1} + \epsilon \quad \text{with} \quad \mathcal{N}(0, 0.007^2) \tag{21}$$

- Discretize the process using both methods, simulate using $T = 10,000$ and compute some moments

	Tauchen			Rouwenhorst		
	$\hat{\mu}_y$	$\hat{\sigma}_y$	$\hat{\rho}$	$\hat{\mu}_y$	$\hat{\sigma}_y$	$\hat{\rho}$
n = 5	0.0009	0.0042	0.9969	-0.001	0.0069	0.9753
n = 9	-0.0013	0.0076	0.9778	0.0003	0.0071	0.9763

Beyond AR(1) processes

- Tauchen method can be extended to VAR(1) processes
  * ARMA(p,q) can be written as VAR(1)
  * VAR(p) can also be written as VAR(1)
- What if the process is not stationary?
  * Typically the case of the income process in quantitative life-cycle models
  * Fella, Galliponi and Pan (2019, Rev. Econ. Dyn.)